The role of time perception in temporal binding:

Impaired temporal resolution in causal sequences

Richard Fereday[b, a], Marc J. Buehner[a], & Simon K. Rushton[a]

[a]School of Psychology, Cardiff University

[b]Dept. Of Psychology, Birmingham City University

Corresponding author:
Richard Fereday
Dept. of Psychology
Birmingham City University
University House
15 Bartholomew Row
Birmingham
B5 5JU
UK
Email: Richard.fereday@bcu.ac.uk
Phone: +44 (0) 121 300 4432

Competing interests

The authors declare they have no competing interests.

Supplementary Materials

The data from these experiments are publically available on the Open Science Framework (on EU-based servers). The data consists of a single Excel file with the aggregated data from each experiment. The data are available here: Fereday, R. (2019, March 7). Temporal binding and internal clocks. Retrieved from osf.io/ar7zk.

Abstract

Causality affects our perception of time; events that appear as causally related are perceived as closer together in time than unrelated events. This effect is known as temporal binding. One potential explanation of this effect is that causality slows an "internal clock" that is used in interval estimation. To explore this hypothesis, we first examined participants' perceived duration of a range of intervals between a causal action and an effect, or between two unrelated events. If (apparent) causality slows the internal clock, then plotting perceived duration against actual duration should reveal a shallower slope in the causality condition (a relative compression of perceived time). This pattern was found. We then examined an interesting corollary: that a slower rate during causal sequences would result in reduced temporal acuity. This is what we found: Duration discrimination thresholds were higher for causal compared to non-causal sequences. These results are compatible with a clock-slowing account of temporal binding. Implications for sensory recalibration accounts of binding are discussed.

Keywords: temporal binding; causal binding; internal clocks; sensory recalibration; time perception

The role of time perception in temporal binding: Impaired temporal resolution in causal sequences

1. Introduction

When a cause (e.g., an action) triggers an outcome, the action is perceived to occur later and its outcome earlier than two unrelated events (Buehner & Humphreys, 2009; Haggard, Clark & Kalogeras, 2002); and the temporal interval separating causally related events is perceived as shorter (Humphreys & Buehner, 2009; Nolden, Haering & Kiesel, 2012). This is known as *temporal binding.* One explanation of the binding effect is based on Bayesian inference (Buehner, 2012; Eagleman & Holcombe, 2002): The brain learns that events that appear as causally related are typically contiguous in time, i.e. the temporal gap between them is zero. When an observer judges a temporal gap, the observer combines sensory information about the duration of the gap, with the information about the most likely gap (based on prior experience) which tends to zero. When combined, the result is an under-estimation of the actual duration.

Haggard, Clark & Kalogeras, (2002) discovered temporal binding using the Libet clock method (Libet, Gleason, Wright & Pearl, 1983), in which participants watch a fast-moving clock hand that completes a rotation every 2560 ms. On one set of baseline trials, participants made a voluntary keypress and were asked to report the clock hand position at the time of keypress. On another set of baseline trials they heard a tone and were asked to report the clock hand position when it occurred. On experimental trials, participants made a voluntary key-press that triggered a tone 250ms later and were asked to report either the clock hand position at the time of the keypress or the time of the tone that followed. Judgment error was the difference between the actual and the reported event (keypress or tone) time. Judgment errors from experimental trials

revealed systematic shifts relative to single-event baseline trials: key presses were subjectively delayed, while their outcomes were perceived to have happened earlier. Put simply, actions and their outcomes appear to attract each other in subjective time.

Since the original Haggard et al. (2002) study, many labs have replicated and extended the temporal binding effect, suggesting, amongst other things, that varying the duration and predictability of the action-outcome interval affects the extent of temporal binding, such that the extent of binding increases for longer intervals (Ruess, Tomaschke & Kiesel, 2018). A different approach to binding has been pursued by Buehner and Humphreys (2009), who used a stimulus anticipation method (SAM). In this study, participants were trained to synchronise their key presses to two tones, t1 and t2. In a causal condition a key press coinciding with t1 triggered the outcome tone, t2. In a non-causal condition participants were still told to press the key to coincide with t1, but in this condition the key press was not causally related to the outcome tone t2; instead, t2 was programmed to follow t1, regardless of whether or not the key was pressed. In other words, in the causal condition, participants were the cause of t2, but in the non-causal condition, the computer was the common cause of both t1 and t2. Buehner & Humphreys found early anticipation (indexed by objectively earlier key presses) of the outcome tone, t2, in causal, relative to non-causal conditions. This study is notable in demonstrating that intentional action and the ability to predict the outcome via association on their own are not sufficient for temporal binding. Instead, it is necessary that action and outcome are causally related for binding to occur. The role of causality in binding is further supported by evidence that an outcome triggered by a non-intentional mechanical device also results in binding (Buehner, 2012): In this preparation, participants pressed a key (self-causal condition) or watched a robot press the same key (machine-causal condition) to produce a subsequent target LED flash; in a

baseline condition, participants watched a signal LED flash followed by the target flash. The key-press → target interval in the causal conditions was identical to the signal-target interval in the baseline condition. In all conditions, participants had to anticipate the target flash by attempting to press a response key simultaneously with it. In line with the causal binding hypothesis, Buehner (2012) found earlier awareness (as indexed by the key-press timings) of the target flash in both the self-causal as well as the mechanical-causal condition, relative to baseline. Taken together, these results indicate that awareness of causality is necessary and sufficient to produce temporal binding, and intentional motor action is neither necessary nor sufficient.

If the perceived timing of events is changed by causal relations then it follows that the perceived gap between events should also change. It does. Studies find that a causal interval is perceived as shorter in time than a non-causal interval, using a variety of methods: interval (magnitude) estimation, in which participants make verbal estimates of the interval (Cravo, Haddad, Claessens & Baldo, 2013; Engbert et al., 2008; Humphreys & Buehner, 2009; Moore, Wegner & Haggard, 2009; Wenke & Haggard, 2009); temporal reproduction, where participants hold down a key for the duration of the experienced interval (Humphreys & Buehner, 2009); and the method of constant stimuli, where a range of variable duration comparison intervals are compared to fixed duration reference intervals (Nolden, Haering & Kiesel, 2012). Thus, binding manifests not only as shifts in event perception, but also as a direct shortening of the experienced time interval.

1.1. Theories of temporal binding

One account of temporal binding implicates the human motor system. A common variant of this approach involves forward models of motor control (Wolpert &

Ghahramani, 2000), which consider processes dedicated to voluntary action. According to this view temporal binding arises from the perceptual system's effort to match sensory feedback to a strongly predicted signal, in line with efficient action-outcome learning (Haggard et al., 2002) and sense of agency (Moore & Obhi, 2012). A variant of this account explains binding as emerging from sensory recalibration (Eagleman & Holcombe, 2002; Parsons, Novich & Eagleman, 2013). For example, when performing finger clicks, the visual, auditory, and tactile information arrives at our brains with different latencies. Yet, our brain compensates for these latencies by realigning the sensory streams to ensure a unitary percept. Eagleman and Holcombe argue that because causal actions usually produce outcomes immediately, the perceptual-motor system realigns input streams to bring action and outcome closer together in time, in an effort to restore unity of perception (cf. the Bayesian Inference account referred to at the beginning of this article). Note that both perspectives only account for binding when the motor system is engaged in the task. Buehner's (2012) result of causal binding following mechanical (robot) action suggests that Bayesian adaptation of event perception operates according to more general principles, namely when the observer is aware of a causal relation between two events.

A different theoretical approach suggests that binding might arise due to changes in time perception (Wenke & Haggard, 2009; Fereday & Buehner, 2017). Because there is no dedicated sense modality for time, temporal information is gleaned from, and affected by, various factors: For example, top-down processes, such as paying greater attention to time and increased arousal, both subjectively expand temporal intervals (Droit-Volet, Fayolle, Lamotte, & Gil, 2013; Tse, Intriligator, Rivest & Cavanagh, 2004); bottom-up processes such as the perceptual organization of a stimulus set affect temporal judgments (filled intervals are judged as longer than unfilled: Wearden, Norton,

Martin & Montford-Bebb, 2007); and modality differences, such that auditory stimuli are judged longer than visual stimuli (Wearden, Edwards, Fakhri & Percival, 1998). Common explanations for these temporal distortions typically refer to internal clock (or pacemaker–accumulator) models (Gibbon, Church & Meck, 1984). These models represent the neural accumulation during a to-be-timed interval with discrete pulses. Consider an example: when a to-be-timed interval begins, a switch closes (L_c), allowing pulses to flow from a pacemaker to an accumulator. The pulses are emitted at rate *r*. When the interval ends, the switch opens (L_o), thereby preventing further pulses from being accumulated. Perceived duration is a function of the number of pulses in the accumulator, such that for any real-time interval, more accumulated pulses lead to a longer perceived duration. Thus, one way that internal clock models can explain temporal distortions is via differences in rate *r*. For example, arousal (such as exercise, or an emotional response, such as fear) increases *r* (Stetson, Fiesta & Eagleman, 2007), leading to subjectively longer intervals than in control conditions. Another way in which internal clock models can account for alterations to subjective time is via switch latencies, such that the switch to begin timing (L_c) is activated later in an experimental than a control condition, leading to shorter estimates in the former.

The signature of a difference in clock rate is a systematically increasing difference between perceived vs physical interval duration as the interval duration increases. Consider a concrete example: Let's imagine that Alice and Bob are listening to the 10s, but due to prior exercise, Bob has higher arousal than Alice. If Alice's internal clock registers 100 pulses during the 10s, Bob's clock might register 110 in the same time. If Alice and Bob listed to a 20s and 30s tone, then Alice's clock would register 200 and 300 pulses, while Bob's would register 220 and 330: The difference between Alice's and Bob's subjectively elapsed time would increase as the to-be-timed

interval increases. Formally, this is assessed by regressing perceived and physical durations and measuring the steepness of the slope: the slower the clock, the shallower the slope. One example of pacemaker rate differences can be found in a study by Wearden et al. (1998), who found a steeper slope for auditory compared to visual stimuli: The auditory – visual judgment difference increased as a function of duration. Other studies, such as those that compare filled to unfilled intervals (Wearden et al., 2007) also find evidence of pacemaker rate differences, as revealed by the same regression procedure. The advantage of conducting regression analyses is that, while different slope coefficients implicate pacemaker rate changes, a difference in intercept coefficients implies different switch latencies between conditions. Thus, regressing subjective over objective durations is an efficient means of disentangling the contribution of switch latency from pacemaker rate differences.

1.2. Evidence for slower pacemaker in temporal binding

Can a slower pacemaker rate explain the contraction of time in temporal binding? In other words, if Alice presses a key to produce a tone 500ms later, or merely listens to two tones separated by 500ms, would her pacemaker rate be slower in the former compared to the latter? There is only limited evidence pertaining to this question. Wenke and Haggard (2009) had participants estimate the duration of the interval between two events: in a voluntary condition, a key press triggered an outcome tone after a delay; in an involuntary condition, participants' fingers were passively moved by a servomotor to press the key, followed by an outcome tone after the same delay. Wenke and Haggard found shorter estimates of intervals triggered by voluntary causal actions, relative to intervals following involuntary actions – the typical binding effect. Importantly, Wenke and Haggard embedded a temporal discrimination task into

the trials: participants received two successive electric shocks to the index and middle fingers respectively. The interval between the shocks varied according to a staircase procedure. Participants were asked to judge whether the shocks were simultaneous or successive. The objective of the temporal discrimination task was to probe for clock slowing in voluntary intervals. A slower pacemaker rate would result in larger inter-pulse intervals, thereby increasing the difficulty of making simultaneity judgments, and hence increasing the temporal discrimination threshold. Wenke and Haggard found just this, with higher thresholds for shocks occurring early in the interval in voluntary trials, and attributed this finding to a slower clock early in voluntary intervals. However, while intervals in voluntary trials were judged as shorter than intervals from involuntary trials, there was no slope effect: The underestimation of voluntary compared to involuntary intervals remained constant as interval size increased, contrary to what one would expect from a slowed pacemaker. We therefore have to be careful to interpret these results as evidence for causality-induced clock-slowing, especially since Tomassini, Gori, Baud-Bovy, Sandini & Morrone (2014) have subsequently demonstrated that action preparation alone leads to effector-specific temporal compression: Participants had to make simultaneity judgments of tactile stimulation delivered to their hands, similarly to Wenke and Haggard. Crucially, in Tomassini et al.'s preparation, stimulation was delivered either to a hand involved in a simple action, or to a stationary hand, and unlike in Wenke & Haggard, the action triggered no further consequence. Nonetheless, Tomassini et al. found higher simultaneity thresholds for stimulation delivered to the moving hand immediately before and after motion onset, relative to stimulation delivered to the stationary hand, or to the moving hand after the motion was completed. Motion onset thus leads to effector-specific, transient temporal compression. Wenke and Haggard's threshold

results could thus be entirely driven by such transient temporal compression, which would also explain why they did not obtain a slope effect for interval judgments. Why would there not have been causality-induced slowing of the pacemaker though? A review of Wenke & Haggard's method reveals that while they contrasted voluntary against involuntary movement, the causal relationship remained constant: Both voluntary and involuntary finger movements caused the subsequent tone. It is thus entirely possible that the results they obtained do in fact not reflect temporal binding at all, but are simply caused by the transient temporal compression following motor preparation.

A study that provides some evidence of clock slowing in temporal binding is Humphreys and Buehner (2009), who asked participants to provide verbal estimates of causal and non-causal intervals for a range of durations from 150ms to 4000ms. A regression of estimates onto actual interval durations found shallower slopes for causal, compared to non-causal intervals: in other words, the amount of under-estimation of causal compared to non-causal intervals grew as a function of objective duration, as predicted by a slower pacemaker rate *r*. This evidence is only tentative, however, because verbal estimates of duration frequently violate the assumption of mean accuracy (Wearden and Lejeune, 2008), which refers to the requirement that mean estimates should vary linearly with actual duration (i.e. duration estimates should increase in line with objective interval durations). A common finding though, is that data conforms to Vierordt's law, whereby short durations are judged as longer, and longer durations are judged as shorter, than actual durations. This would lead to shallower slopes than those derived from other methods (e.g., the method of constant stimuli; see description of Nolden et al., 2012, below). A further reason to be sceptical of verbal estimation is that Matthews (2011) found differences in slopes between

conditions where a pacemaker increase/decrease cannot have occurred. Verbal estimation, therefore, is not sufficiently reliable a method with which to investigate differences in slope coefficients.

1.3. The current study

One reason why interval estimation may violate the assumptions of internal clock models is because duration judgments are based on an internal reference stimulus, which is prone to distortion. A better option is to use discrimination methods, which present the reference stimulus during the experiment. Nolden et al. (2012) replicated the binding effect using the method of constant stimuli. Participants judged whether a series of comparison durations were shorter or longer than a fixed reference interval. For causal and non-causal conditions, the authors estimated the point of subjective equality (PSE), which refers to the duration of the comparison interval that is perceived as the same (50% of the time its perceived as longer, 50% of the time its perceived as shorter) as the reference interval. Their results revealed shorter PSEs in causal versus non-causal conditions, for both reference intervals tested (250 and 600 ms). However, there were too few reference durations employed to obtain a reliable estimate of clock speed using the regression method. Therefore, in the present study we adopted a psychophysical procedure similar to that of Nolden et al., and combined it with the regression method of previous research (e.g., Wearden et al., 1998). We regressed perceived intervals (causal and non-causal) onto actual durations to determine whether there would be reliable differences in slope, as predicted by a slowing of *r*.

In addition to pacemaker rate in temporal binding, we were also interested in the impact of a slower clock, i.e., the corollary that temporal resolution is necessarily impaired if the pacemaker operates at a slower rate (cf. Stetson et al., 2007, and Wenke & Haggard, 2009). Specifically, if the contraction of time in temporal binding reflects a decrease in the rate of a specific clock system, then the impaired temporal resolution that follows generates an interesting prediction: The threshold required to temporally discriminate causal intervals should be higher than for non-causal intervals. Importantly, this would not be the case if binding were rooted in shifts in the events that delineate the interval (or by a difference in switch latencies). Only a subjective distortion of time *during* the interval could explain temporal discrimination differences between causal and non-causal intervals.

We report four experiments that investigate clock slowing in temporal binding. In Experiments 1 and 2, there were two conditions, causal and non-causal, and participants experienced two intervals per trial in each: on causal trials, one of the intervals was between a key press and visual flash, while on non-causal trials, one of the intervals was delineated by two visual flashes. In both conditions a single temporally extended stimulus served as the comparison interval. Participants judged whether the single stimulus interval was shorter or longer than the key press-flash interval (causal trials) or the flash-flash interval (non-causal trials). Based on responses from a range of comparison (extended stimuli) interval durations we computed the PSE as a measure of perceived interval duration. We expected shorter causal versus non-causal PSEs, due to temporal binding. Furthermore, a regression of PSEs onto actual duration would reveal slope differences if pacemaker rates vary between causal and non-causal intervals.

In Experiments 3 and 4, we corroborated the findings of Experiments 1 and 2, that revealed shallower slopes in causal conditions. Participants discriminated between

two causal intervals in one condition, and two non-causal intervals in another. In both conditions, participants judged whether one interval was shorter or longer than the other (Experiment 3) or judged which interval was the longest (Experiment 4). We computed the just-noticeable-difference (JND) as a measure of temporal discrimination, which is the minimum duration necessary to discriminate the duration of two intervals. If temporal causal binding is effected via a slower clock rate, then JNDs should be higher for causal intervals, due to the poorer temporal resolution.

The experiments in this paper operationalised the causal vs. non-causal distinction by comparing action-event to event-event intervals (e.g. key-press → flash vs. flash – flash). The bulk of previous research in temporal binding was concerned with a distinction between active causal actions and passively induced non-causal involuntary actions (e.g. Haggard et al., 2002; Wohlschläger et al., 2003; Wenke & Haggard, 2009), to isolate the active, intentional aspect of causal actions from the mere motor component. However, subsequent research has found that a) the active, intentional aspect of the action on its own is insufficient to result in binding – the action has to be causal (Buehner & Humphreys, 2009); b) passively induced actions that cause a subsequent outcome still result in binding, despite the absence of intentional action planning (Buehner, 2015); c) observed causal action-outcome intervals lead to comparable temporal binding as self-executed causal action-outcome intervals (Poonian & Cunnington, 2013); d) cause-effect intervals where the cause is a mechanical action rather than a human motor action, are subject to temporal binding, albeit reduced compared to motor-action causality (Buehner, 2012; Shiloh, White & Buehner, 2017). As mentioned earlier, causality is necessary and sufficient to produce temporal binding, and intentional motor action, while possibly exerting a boost on temporal binding, is neither necessary nor sufficient. We decided on implementing the causal vs non-causal

distinction by comparing temporal intervals between a motor action (key press) that triggers an outcome to intervals between two unrelated events, because this is by far the most efficient and economical way to study temporal binding. We discuss potential limitations of this approach in the General Discussion.

2. Experiment 1

We conducted the following experiment in order to (1) determine whether clock slowing underpins temporal binding; (2) explore the involvement of switch latencies; and (3) determine whether timing sensitivity is constant or variable, in cause-effect intervals. To these ends, we conducted a variant of Wearden et al. (1998). In their study, interval estimates were regressed onto actual durations and the resultant coefficients analysed to disentangle the contribution of pacemaker rate and switch latencies. In the current experiment though, we replaced interval estimation with a discrimination procedure. In causal and noncausal conditions participants experienced reference and comparison intervals. The reference interval was delineated by two events: a key press and visual flash (causal trials) or two visual flashes (noncausal trials). The comparison interval was a single temporally extended visual stimulus in both conditions. The sequence of reference and comparison intervals was manipulated with a between-subjects factor: either the reference interval first and comparison, or the reverse. In both levels of the between-subjects variable, participants compared the duration of the comparison to the reference interval. The PSEs were then estimated and regressed onto reference durations to analyse the coefficients. If temporal causal binding is effected via the slowing of an internal clock, we should find shallower slopes for causal versus noncausal conditions in addition to lower PSEs.

2.1. Method

2.1.1. Participants

Thirty-five students of Cardiff University (33 female, 2 male, M_{age} = 19.76, age range: 18 - 29) participated in Experiment 1. Participants received course credit or £5 payment.

2.1.2. Apparatus and stimuli

The experiment was implemented in Psychopy (Peirce, 2007) on cathode ray tube (CRT) monitors with resolution of 1280 x 1024 and refresh rate of 120 Hz connected to Apple Mac Minis. Both conditions (causal, noncausal) featured two intervals. A reference interval was marked by two events: a key press and the appearance of a white square in causal trials, and two white squares in noncausal trials (all white squares were 200 pixels^2 and displayed for 50 ms). The second comparison interval was a single visual event (a black circle; radius 70 pixels). The duration of reference interval was fixed for a block, while the duration of the comparison interval varied from trial to trial, as determined by a staircase procedure. A white fixation cross (60 pixels^2) was displayed throughout each trial. All stimuli were presented centrally on the screen. Figure 1 illustrates the trial structure.

2.1.3. Design and procedure

We employed two factors: Trial Type (causal and noncausal) and Reference Duration (5 durations, 200ms steps: 200, 400, 600, 800, 1000 ms). The dependent variable was participants' PSE as derived from a series of trials controlled by a staircase procedure. Participants completed 5 blocks of causal and noncausal trials (one block per Reference Duration x Trial Type combination). The number of trials within a block varied due to the staircase procedure employed – each block ended when minimum reversal and trial number criteria were satisfied as explained below , with each block

comprising around 30 trials. This resulted in around 300 trials total, with an additional 12 practice trials, prior to the experiment, for causal and noncausal blocks.

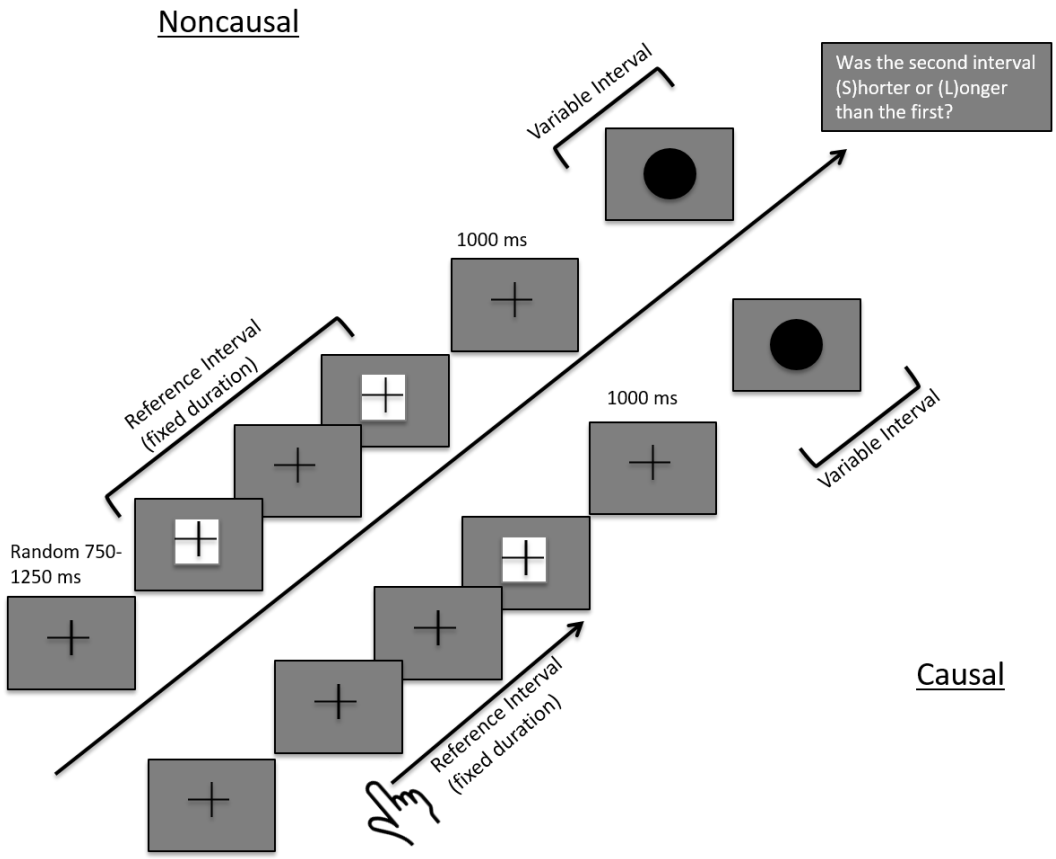

Figure 1. Representative trial structure of Experiments 1 and 2. Participants were presented with two intervals in each condition. Half of the participants experienced the following: the duration of the first interval was fixed for a block, while the second was variable. In causal trials, a key press triggered a visual flash after a delay (interval 1), followed by an extended temporal event (black circle: interval 2). In noncausal trials, two visual flashes delineated interval 1, followed by the black circle (interval 2). Participants were asked whether the second interval was shorter or longer than the first. The remaining half were presented with the variable interval first, then the reference interval. They were then asked whether the first interval was shorter or longer than the second.

Causal trials began with the display of a fixation cross on screen. Participants were told to press a key at a time of their choosing, which triggered a visual flash (white square as per above) after a fixed duration (200, 400, 600, 800, 1000 ms). After an ISI of 1000 ms, the black circle was displayed for a duration that was varied from trial to trial by a staircase algorithm. Participants were then asked whether the duration of the black circle was shorter or longer than the key press - flash interval. After responding, the screen blanked for a random duration (1500 - 2300 ms) before the display of the fixation cross signalled the next trial.

Noncausal trials followed a similar procedure. The trial began with the display of a fixation cross, presented for a random duration (750 - 1250 ms). A visual flash (white square) marked the start of the first interval, followed by another flash of a white square after a fixed interval (200, 400, 600, 800, 1000 ms). After an ISI of 1000 ms, the second interval (black circle) was displayed for a duration (variable from trial to trial). Participants were then asked whether the second (variable) interval was shorter or longer than the first (fixed) interval. Participants responded by pressing the S or L keys, respectively. The screen then blanked for a random duration (1500 - 2300 ms) before the fixation cross signalled the beginning of the next trial.

The experiment began with either a causal or noncausal block (first completing practice trials: see below), thereafter alternating between them. The duration of the fixed interval was selected at random from one of the five reference duration intervals, and remained fixed for each block. The variable interval was controlled by a staircase procedure, using a Kesten stochastic approximation algorithm (Kesten, 1958; Treutwein, 1995). This allows for fast and reliable convergence onto the point where participants are 50% correct. We used two randomly interleaved staircases, one

ascending one descending, each with target convergence thresholds of 0.5, with each block ending after a minimum of 4 reversals and 15 trials per staircase. The first trials of each staircase began with durations of 0.5 and 1.5 times the fixed duration, i.e., for a fixed interval of 400 ms, the variable interval would begin at durations of 200 and 600 ms for each respective staircase. Additionally, the first trial of each staircase was classed as a control trial, with the first and second trials of each staircase set to the same duration, regardless of response. This was to ensure that incorrect or mistaken responses early in each block did not adversely affect the direction of each staircase, e.g., a mistaken response would alter the direction of the staircase and reduce the likelihood of convergence within the scheduled number of trials.

Participants first completed practice blocks (causal and noncausal), which contained a fixed reference duration of 650 ms. These were shorter than the main experimental blocks and were programmed to end after a minimum of two reversals and six trials per staircase. Participants then alternated between causal and noncausal blocks of the main experiment. Throughout the experiment, participants were given a 2-minute break between blocks. The study was conducted across two sessions, with one session per day. Participants began with a causal block in session one and noncausal in session two.

We reversed the order of fixed (reference) and variable intervals in a between-subjects manner. For half of participants, the comparison interval was presented first, and the reference interval second. For the other half, the reverse applied (we operationalised this as the factor Interval Order [reference first, variable first]). The fixation cross disappeared during the ISI period, so that participants would know when to press the key to trigger the flash. Participants were still asked whether the comparison interval was shorter or longer than the reference interval.

2.2. Results

PSE analysis[1]. Prior to statistical analysis, we removed the first (control) trial of each staircase, before fitting cumulative Gaussian curves, following the generalized linear model (GLM) procedure in R (Knoblauch & Maloney, 2012). We then calculated the PSEs and screened the data for normality. Participants whose PSEs deviated from the mean by more than 3 standard deviations (in any Trial Type x Reference Duration combination), were not entered into the analysis. This led to ten participants being removed[2]. Figure 2 shows that causal PSEs are shorter than non-causal for all levels of Reference Duration. Figure 2 also shows a linear relationship between PSEs and actual durations, indicating that participants successfully distinguished between reference durations. ANOVA supports these observations, with a significant effect of Trial Type, $F(1, 24) = 23.31$, $p < .001$, partial $\eta^2 = .49$, Reference Duration, $F(4, 94) = 53.13$, $p < .001$, partial $\eta^2 = .69$, and a Trial Type x Reference Duration interaction, $F(4, 96) = 3.98$, $p < .01$, partial $\eta^2 = .14$. This also supports the observation that the causal vs non-causal difference increased with Reference Duration. Post hoc comparisons find significant differences between all levels of Trial Type (200 ms: $p < .05$; 400, 800 ms: $p < .01$; 600, 1000 ms: $p < .001$). Whilst we did not find an effect of Interval Order, $F(1, 24) = .03$, $p = .86$, partial $\eta^2 < .01$, or a Trial Type x Interval Order interaction, $F(1, 24) = 2.68$, $p = .12$, partial $\eta^2 = .10$, the Reference Duration x Interval Order interaction was significant, $F(4, 96) = 4.16$, $p < .01$, partial $\eta^2 = .15$ (See Online Supplement and Figures S3 and S5 for details). The Trial Type x Reference Duration x Interval Order interaction was not significant, however, $F(4, 96) = .59$, $p = .67$, partial $\eta^2 = .02$.

Regression analyses[3]. To ascertain whether switch latencies or pacemaker slowing can explain the data, we conducted further analyses. Perceived durations were regressed onto actual durations for each participant. We then performed mixed ANOVAs on the resultant intercept and slope coefficients, with Trial Type as a within-subjects factor and Interval Order (2 levels [reference first, variable first]) as a between-subjects factor. The results are as follows:

Intercepts. The analysis showed no significant effect of Trial Type, $F(1, 22) = 2.02$, $p = .17$, partial $\eta^2 = .08$, nor a Trial Type x Interval Order interaction, $F(1, 22) = 1.13$, $p = .30$, partial $\eta^2 = .05$. However, we did find a significant effect of Interval Order, $F(1, 22) = 32.57$, $p < .001$, partial $\eta^2 = .60$ (See Online Supplement and Figure S1 for details).

Slopes. The analysis showed a significant effect of Trial Type, $F(1, 22) = 10.10$, $p < .01$, partial $\eta^2 = .32$, which crucially, supports our hypothesis of a slower clock in causal conditions. The ANOVA also revealed a main effect of Interval Order, $F(1, 22) = 12.92$, $p < .01$, partial $\eta^2 = .37$, but no Trial Type x Interval Order interaction, $F(1, 22) = 2.64$, $p = .12$, partial $\eta^2 = .12$ (See Online Supplement and Figure S2 for details).

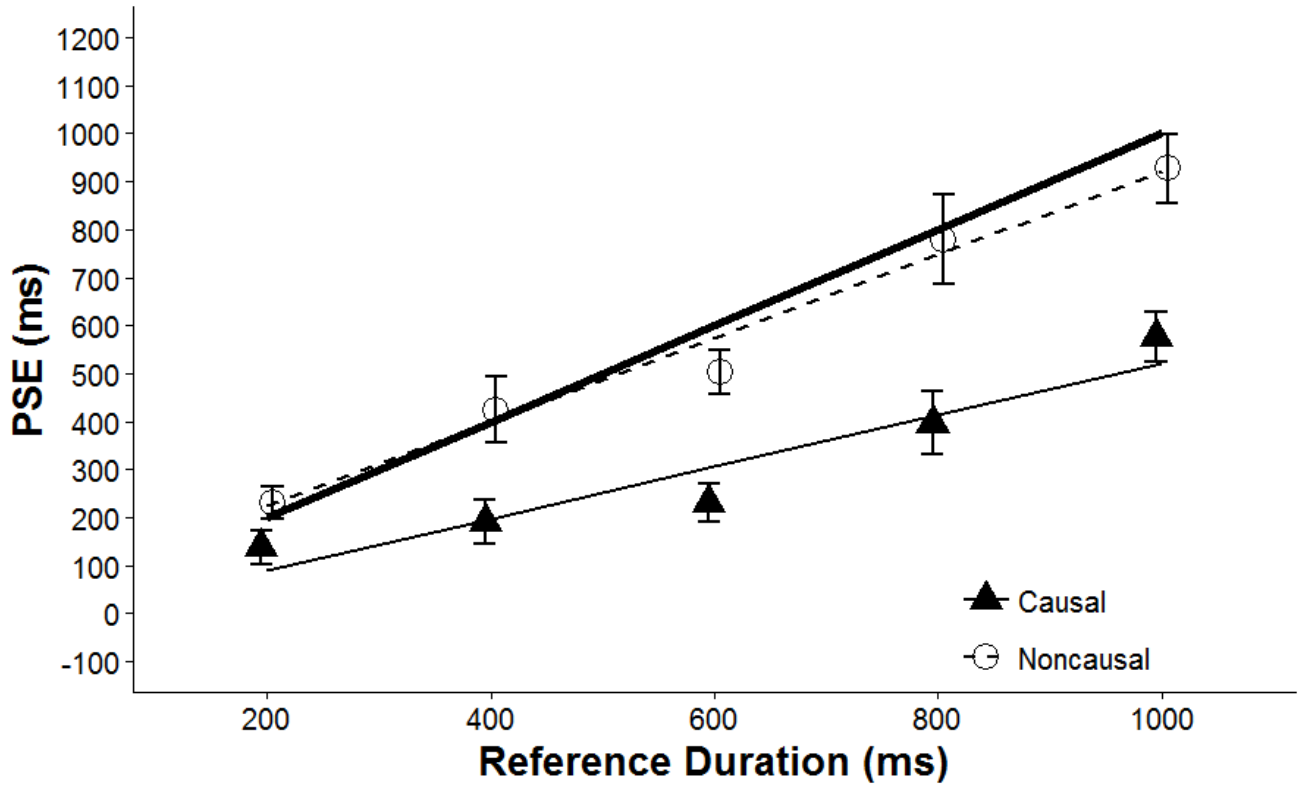

Figure 2. Results of Experiments 1. Causal PSEs are shorter than Non-Causal, thus replicating the temporal binding effect. The bold line represents the veridical responses; this shows that causal PSEs are underestimated across all Standard Durations. Error bars show standard error.

2.3 Discussion

In the current experiment we replicated the temporal binding effect using a temporal discrimination procedure. Unsurprisingly, we found that PSEs increased linearly with reference duration in both conditions, indicating that participants successfully distinguished between durations. Also, PSEs were lower than objective reference durations in causal conditions only; noncausal PSEs were largely veridical. More importantly though, we found lower PSEs in causal versus noncausal conditions, in line with temporal binding.

However, our main goal was to identify whether temporal binding implicates changes to pacemaker speed or switch latency. To this end, we calculated regression coefficients on a participant-by-participant basis, and then compared mean slope and intercept values between conditions. Slope differences would be evidence for different pacemaker speeds; specifically, a shallower slope for causal compared to non-causal intervals would indicate that the relative underestimation of causal intervals is driven by a slower pacemaker during causal episodes. Intercept differences, on the other hand, would be evidence for delayed closing of the switch; specifically, a smaller intercept for causal compared to non-causal intervals would indicate that relative underestimation of causal intervals is driven by a delayed closing of the switch at the start and/or the end of causal episodes.

We found no evidence for a switch-latency effect, finding no difference between causal and noncausal intercepts. Instead, the results of the slope analysis finds shallower slopes in causal than noncausal conditions. This supports a clock slowing account of temporal binding. However, to substantiate the current findings, we conducted a further experiment with more Reference Durations.

3. Experiment 2

The second experiment was a conceptual replication of Experiment 1.. We increased the number of reference durations from five to ten. We maintained the same range of durations but increased the resolution, using steps of 100 ms (i.e., 100, 200, 300, 400 . . . 1000 ms). In other respects, the procedure was identical.

3.1. Method

3.1.1. Participants

Thirty-four students of Cardiff University (29 female, 5 male, M_{age} = 20, age range: 18 - 25) participated in Experiment 2 (we applied the same exclusionary criteria as in Experiment 1: see footnote 1). Participants received course credit or £10 payment.

3.1.2. Apparatus and stimuli

This remained the same as Experiment 1.

3.1.3. Design and procedure

The procedure was identical to Experiment 1 with the exception that we used more reference durations (100, 200, 300, 400 . . . 1000 ms). We also manipulated the order of reference and comparison intervals along the lines of Experiment 1, using a between-subjects variable. Participants completed around 300 trials total, with an additional 12 practice trials, prior to the experiment, for causal and noncausal blocks.

3.2 Results

PSE analysis. We followed the same curve fitting and exclusion procedures as in the previous experiments. This resulted in the removal of 9 participants. Figure 3 shows that all PSEs increase linearly with reference durations, implying that participants distinguished between reference durations. Importantly, causal durations are perceived as shorter than noncausal, in line with the results of the previous experiment. Statistical analyses support these observations, with ANOVA showing a significant effect of Trial Type, $F(1,22) = 26.63$, $p < .001$, partial $\eta^2 = .55$, Reference Duration, $F(9,198) = 50.67$, $p < .001$, partial $\eta^2 = .70$, and a Trial Type x Reference Duration interaction, $F(9, 198) = 3.34$, $p < .01$,partial $\eta^2 = .13$. This interaction supports our observation that the causal-noncausal difference increases linearly with Reference Duration. Post hoc tests find significant differences between Trial Type for most levels of Reference Duration (100, 300, 400, 600, 800, 900, 1000 ms: $p < .01$; 200, 500 ms: $p < .001$; 700 ms: $p = .11$). There

was no effect of Interval Order, $F(1,22) = .04$, $p = .84$, partial $\eta^2 < .01$, nor a Trial Type x Interval Order interaction, $F(1,22) = 2.03$, $p = .17$, partial $\eta^2 = .09$. However, we did find significant interactions between Reference Duration x Interval Order, $F(9,198) = 5.00$, $p < .001$, partial $\eta^2 = .19$, and Trial Type x Reference Duration x Interval Order, $F(9,198) = 1.91$, $p = .05$, partial $\eta^2 = .08$. Illustrations and interpretations for these interactions are provided in the Online Supplement and Figures S4 and S6.

Regression analyses. As in the analyses for Experiments 1 and 2, we conducted individual regressions to further probe switch latency and clock slowing effects. We then performed mixed ANOVAs on the intercept and slope coefficients, with Trial Type as a within-subjects factor and Interval Order as a between-subjects factor. The results are as follows:

Intercepts. The analysis showed no significant effect of Trial Type, $F(1, 24) = 1.54$, $p = .23$, partial $\eta^2 = .06$, nor a Trial Type x Interval Order interaction, $F(1, 24) = 1.64$, $p = .21$, partial $\eta^2 = .06$. However, we did find a significant effect of Interval Order, $F(1, 24) = 11.14$, $p < .01$, partial $\eta^2 = .32$ (See Online Supplement and Figure S1 for details).

Slopes. The analysis showed a significant effect of Trial Type, $F(1, 24) = 18.84$, $p < .001$, partial $\eta^2 = .44$, and Interval Order, $F(1, 24) = 13.49$, $p < .01$, partial $\eta^2 = .36$. The former is line with Experiment 1, and supports our clock-slowing account of temporal binding (See Online Supplement and Figure S2 for details). We found no Trial Type x Interval Order interaction, $F(1, 24) = .22$, $p = .64$, partial $\eta^2 = .01$.

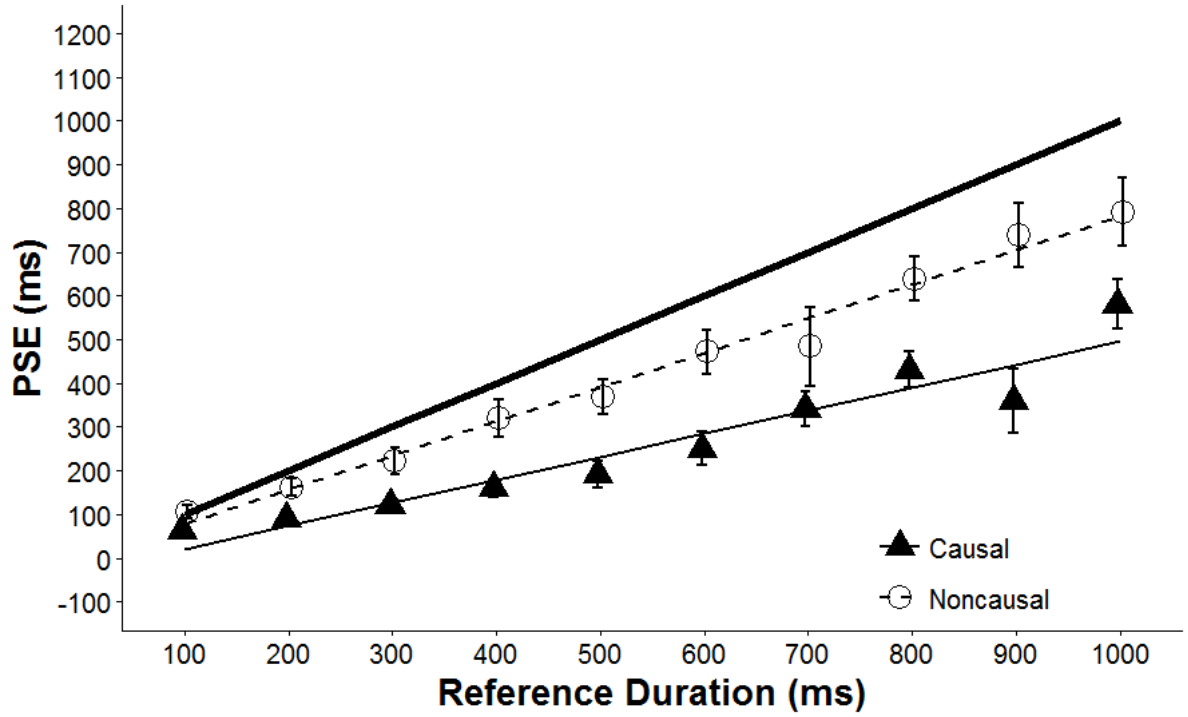

Figure 3. Results of Experiments 2. Causal PSEs are shorter than Noncausal, complementing the results of Experiment 1. The bold line represents the veridical responses; this shows that causal and noncausal PSEs are underestimated across all Standard Durations. Error bars show standard error.

3.3 Discussion

The results of the current experiment replicate those of Experiment 1. PSEs increased linearly in causal and noncausal conditions, which indicates that participants discriminated reference durations. Interestingly, PSEs were underestimated relative to veridical in causal and noncausal conditions. This is in contrast to the previous experiment, where noncausal PSEs were veridical. The crucial finding, however, is that causal PSEs were lower than noncausal, thus supporting the temporal binding effect.

In order to determine whether binding is effected via switch latencies or clock slowing, we examined the regression coefficients using the same procedure as Experiment 1. We found no evidence for differences in switch latencies, with no

significant difference between causal and noncausal intercepts. Importantly, we found a significant difference between slopes, further supporting a slower clock in temporal binding.

4. Experiment 3

The two experiments reported thus far have provided evidence in support of the clock-slowing hypothesis. A corollary of a slower clock during causal episodes is that temporal discrimination should be impaired, relative to non-causal intervals (see Figure 4). We thus tested the prediction that a slower clock rate in causal intervals would affect temporal acuity. Specifically, when discriminating two causal intervals (in one condition) and two noncausal intervals (in another) fewer pulses in causal intervals should lead to poorer temporal resolution, which in turn would result in poorer discrimination acuity.

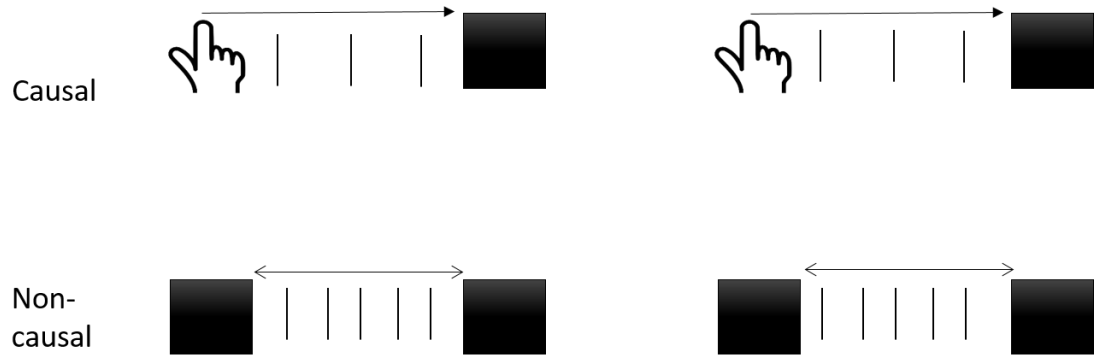

Figure 4. Schematic showing hypothesised internal clock pulses in causal and noncausal intervals (Experiments 3 and 4). The pulse rate *within* conditions is stable, but varies *between* conditions: Fewer pulses in causal intervals, as effected by a slower pacemaker, would result in poorer temporal discrimination when comparing two causal intervals to one another relative to comparing two noncausal intervals.

To investigate this prediction, we conducted a temporal discrimination task, which measured discrimination thresholds in causal and noncausal scenarios. In causal trials participants had to discriminate between two intervals delineated by a key press and visual flash. In noncausal trials, participants were asked to discriminate two intervals delineated by two visual flashes. The duration of the first interval in both conditions was fixed for a block, while the second varied from trial to trial. Participants were asked to judge whether the second (variable) interval was shorter or longer in duration than the first (fixed) interval. We then computed the discrimination threshold (JND) for a range of intervals. Larger JNDs in causal trials would be compatible with a slower pacemaker rate during causal intervals, relative to noncausal.

Note that PSEs for comparing causal to causal and non-causal to non-causal intervals should not differ. In contrast to Experiments 1 and 2, where participants compared causal vs non-causal intervals to a temporally extended reference duration, Experiments 3 and 4 deployed comparisons of like intervals. Differences in pacemaker rate would therefore only influence temporal acuity, but not subjective equality. Consider comparing the weights of two rocks, once with high-precision scales and once with low-precision scales. The minimum difference in weight that is required to identify one rock as heavier than the other will be smaller in the former compared to the latter case. At the same, though, accuracy (but not precision) of identifying two rocks of equal weight would be unaffected by scale precision.

4.1. Method

4.1.1. Participants

Fourteen students of Cardiff University (13 female, M_{age} = 23.43, age range: 21 - 42) participated in exchange for course credit or £4 payment.

4.1.2. Apparatus and stimuli

The experiment was implemented in Psychopy (Pierce, 2007) on CRT monitors, with resolution of 1280 x 1024 and refresh rate of 120 Hz connected to Apple Mac computers. In both conditions (causal and noncausal) participants were presented with two intervals. In causal trials both intervals were between a key press and a visual flash (white square against a grey background, 200 pixels^2), while two visual flashes marked the intervals in noncausal trials. All visual flash stimuli were 50 ms in duration. In causal and noncausal trials the first interval was of a fixed duration, while the second always varied in duration from trial to trial according to a staircase procedure. A fixation cross (60 pixels^2) appeared on screen before each interval and was displayed until the interval finished. All stimuli were presented centrally on screen.

4.1.3. Design and procedure

Trial Type (causal, noncausal) and Reference Duration (250, 450, 650 and 850 ms) varied as within-subjects factors. The dependent variable was derived from the participants' discrimination judgments. We operationalized discrimination as the just noticeable difference (JND), which is half of the difference between 0.75 and 0.25 responding probability from a fitted psychometric function. Participants alternated between causal and noncausal experimental blocks, with one block of each reference duration, comprising roughly 50 trials per block (two interleaved staircases, each ending after a minimum of 25 trials and 4 reversals. See specifics below).

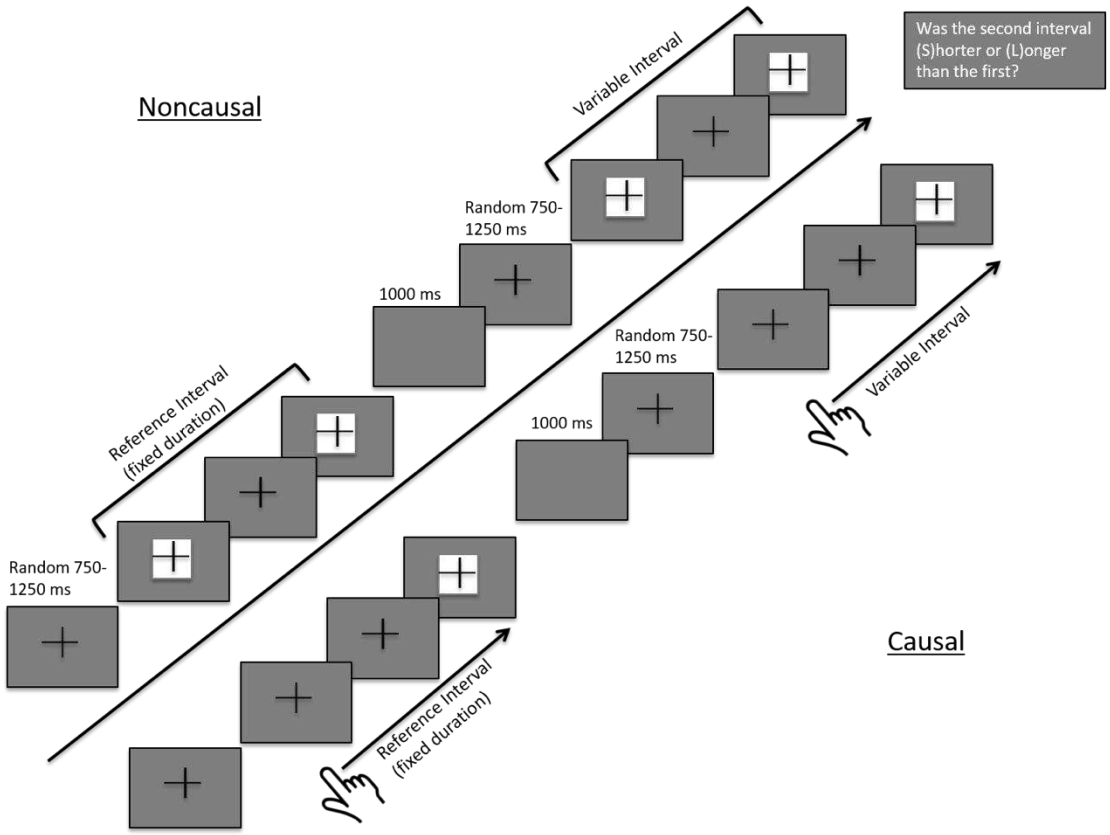

Figure 5. Trial structure of Experiment 3. Participants were presented with two intervals in each condition. The duration of the first interval was fixed for a block, while the second was variable. In causal trials, a key press triggered a visual flash after an interval; in noncausal trials, two visual flashes delineated an interval. Participants were asked whether the second interval was shorter or longer than the first.

Figure 5 shows the trial structure of the experiment. Causal trials began with the display of a fixation cross. Participants then made a voluntary key press, which triggered a visual flash after a fixed interval (250, 450, 650 or 850 ms). During this interval the fixation cross remained on screen, but disappeared immediately following the flash. Participants then experienced an inter-stimulus interval (ISI) of 1000 ms before the fixation cross was again displayed. As in the first interval, participants made a voluntary key press that triggered a visual flash after a variable interval (length

determined by the staircase procedure). The fixation cross again disappeared immediately following the flash. Participants were then asked whether the second (variable) interval was shorter or longer than the first (fixed) interval by pressing the S or L keys, respectively. The screen then blanked for a random duration between 1500 - 2300 ms before the next trial began.

Noncausal trials followed a similar procedure. Trials began with the display of a fixation cross, for a random time (750 - 1250 ms). A visual flash then marked the beginning of the interval, followed by another visual flash after a fixed duration (250, 450, 650 or 850 ms). After an ISI of 1000 ms the fixation again appeared (random duration: 750 - 1250 ms,) before the second interval began, which was delineated by two visual flashes. The fixation cross remained on screen throughout each interval and disappeared for the ISI period. As in causal trials, participants were asked to press the S or L keys if they perceived the second interval to be shorter or longer than the first.

The duration of the fixed interval was selected at random from one of the four Reference Duration intervals, and remained fixed for each block. The variable interval was controlled by a staircase procedure, using the same Kesten stochastic algorithm as in Experiments 1 - 4 (Kesten, 1958; Treutwein, 1995). However, the convergence thresholds of each staircase were 0.25 and 0.75, respectively, and each block ended after a minimum of 4 reversals and 25 trials per staircase. In other respects, the procedure remained the same as the previous experiments.

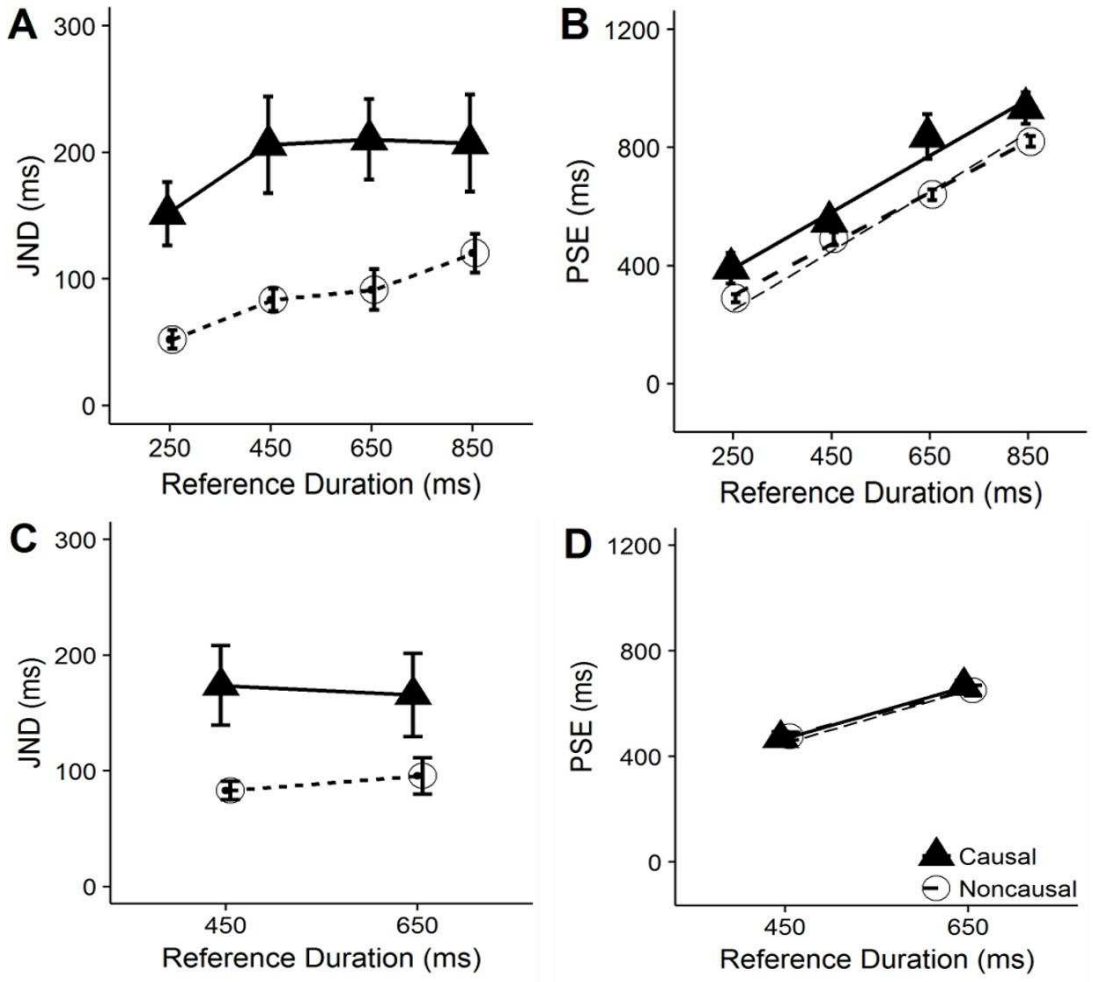

Figure 6. Results of Experiments 3 (upper row: A & B) and 4 (lower row: C & D). (Left panels: A & C): JNDs are higher in causal conditions than noncausal, showing that participants found causal intervals harder to discriminate in both experiments. (Right panels: B & D): In Experiment 3, causal PSEs are overestimated, with respect to the veridical responses (shown by a lighter/shorter dashed line). Error bars represent standard error.

Participants first completed practice blocks, with a fixed reference duration of 600 ms. These were shorter than the main experimental blocks and were programmed to end after a minimum of 2 reversals and 6 trials per staircase. Participants began with either causal or noncausal blocks, the order of which was randomly determined. Participants completed one block of each (causal or noncausal), before moving on to the main experimental blocks. Throughout the experiment, participants were given a 2-minute break between blocks.

4.2. Results

JND analysis. We first applied the exclusion criteria from Experiments 1 and 2, but no participants merited removal. Control trials (trial 1 of each staircase) were removed before fitting psychometric functions (cumulative Gaussian). Figure 6(A) shows mean JNDs for Trial Type, plotted against Reference Duration. JNDs are clearly higher for the causal condition across all reference durations, indicating that participants were less precise in discriminating causal intervals compared to noncausal. A 2 x 4 (Trial Type [causal, noncausal] x Reference Duration [250, 450, 650,850]) ANOVA corroborated these impressions. We observed a main effect of Trial Type, $F(1,13) = 25.06$, $p < .001$, partial $\eta^2 = .67$, and Reference Duration, $F(3,39) = 3.79$, $p < .05$, partial $\eta^2 = .23$. However, the Trial Type X Reference Duration interaction was not significant, $F(3,39) = .42$, $p = .74$, partial $\eta^2 = .03$.

PSE analysis. Because this experiment presented participants with two matched intervals (i.e. two causal intervals in a row), one would of course expect PSEs to be veridical. However, the reference intervals were always presented before the variable duration intervals, creating scope for order effects. For example, a common finding in time perception is the temporal order effect (TOE), where the second interval is judged as longer than the first, when the intervals are presented in series (Schab & Crowder,

1988). Although, to our knowledge, no studies have examined the effect of stimulus order on discrimination judgments (as opposed to duration judgments), we nevertheless wanted to ascertain that order effects are not driving our results. To examine whether order effects were present, we inspected causal and noncausal PSEs; because the experiment involved comparing like-for-like (i.e.,causal v causal intervals, noncausal versus noncausal intervals), the PSEs ought to be (a) veridical (i.e., in line with actual durations); and (b) should not differ between Trial Type conditions. Figure 6 (B) shows the PSEs for each level of Reference Duration. Visual inspection shows a near perfect linear relation between noncausal PSEs and actual stimulus duration. In contrast, causal PSEs appear consistently longer than veridical for all reference durations, suggesting a TOE (the second interval required a longer real-time duration in order to be judged as equal in duration to the first). In other words, the second interval was actually perceived as *shorter* than the first. This observation is corroborated by a 2 x 4 within subjects ANOVA with factors Trial Type and Reference Duration: In addition to the main effect of Reference Duration, $F(3,39) = 182.91$, $p < .001$, *partial* $\eta^2 = .93$, the main effect of Trial Type was also significant, $F(1,13) = 7.80$, $p = .015$, *partial* $\eta^2 = .38$. The interaction was not significant, $F(3,39) = 1.67$, $p = .209$, *partial* $\eta^2 = .11$.

4.3. Discussion

As would expected from the Weber-Fechner Law, JNDs scaled with reference duration. More importantly and in line with the hypothesis of a slower pacemaker in temporal binding, we also found that causal JNDs were significantly larger than noncausal, for all reference durations tested. An unexpected finding, though, concerns the PSEs. While PSEs in noncausal conditions were generally veridical, causal PSEs were larger. This indicates that the second (comparison) causal interval was perceived as shorter than the first (reference) causal interval. In other words, our data exhibit an

order effect, which might (at least, in part) have contributed to our JND results. To avoid conflating order effects with those induced by causality, and to provide more compelling evidence for causality-induced discrimination differences, we conducted one further experiment.

5. Experiment 4

This experiment differed to the previous one only in terms of the sequence of reference and comparison intervals in each trial. We randomised the temporal position of the reference and comparison intervals and asked participants to judge which interval was the longest. Additionally, we used only two reference interval durations. In other respects, the procedure remained the same as in Experiment 3.

5.1 Method

5.1.1. Participants

Six students of Cardiff University (5 female, 1 male, M_{age} = 26.83, age range: 22 - 33) participated in exchange for £2 payment or course credit.

5.1.2 Apparatus and stimuli

This remained the same as the previous experiment.

5.1.3. Design and procedure

This remained the same as the previous experiment with a few exceptions. We randomised the sequential position of the reference and variable intervals, so that the reference interval could be presented either first or second in a particular trial. Also, we used two standard durations (450, 650 ms) instead of four. Participants were asked

which interval was longest, and responded by pressing either the left or right arrow key to correspond to the first or second interval, respectively. In all other respects the procedure remained the same as above.

5.2. Results

JND analysis. We followed the exclusion criteria used in the previous experiments in this paper, with no participants meriting removal. As before, we removed the first trial of each staircase before fitting cumulative Gaussian functions. Mean JNDs are shown in Figure 6 (C). Consistent with the previous results, JNDs are higher in causal than noncausal conditions. ANOVA supports these results, with a main effect of Trial Type, $F(1,5) = 11.19$, $p < .05$, partial $\eta^2 = .69$. In contrast to the previous experiment we found no effect of Reference Duration, $F(1,5) = .02$, $p = .90$, partial $\eta^2 < .01$. The Trial Type x Reference Duration interaction again was not significant, $F(1,5) = .12$, $p = .74$, partial $\eta^2 = .02$.

PSE analysis. A final and important observation concerns order effects. Given that the intervals in the trial sequence were random (i.e., fixed duration intervals could either be first or second in a particular trial), we expected PSEs of both conditions to be veridical. Indeed, this is what we find, as Figure 6 (D) shows. The main effect of Duration was significant, $F(1,5) = 67.36$, $p < .001$, partial $\eta^2 = .93$, but the main effect of Trial Type was not, $F(1,5) = 0.02$, $p = .881$, partial $\eta^2 < .01$. Thus, the finding that JNDs in causal trials are higher than non-causal, cannot be attributed merely to order effects, and instead, is likely due to poorer temporal resolution, as would be expected from a slower pacemaker rate.

5.3. Discussion

The results of the present experiment complement those of Experiment 3. We found that causal JNDs were significantly larger than non-causal. We are reassured to find that randomising the order of reference and comparison intervals eliminated the order effects on PSE we observed for causal intervals Experiment 3, whereby the comparison interval was perceived as shorter than the reference. Instead, PSEs were near veridical and identical in both the causal and non-causal condition. This assures us that the key result of threshold differences differences found in Experiment 3 (and replicated in Experiment 4) could not have been driven by the artificial distortions associated with order effects of reference vs. variable intervals.

Can the results be explained by differences in switch latencies? If the start/stop latencies are related to the interval to-be-timed, such that the latencies are a fixed ratio of the interval, then both intervals would be perceived as equal. In other words, comparing like-for-like intervals would not result in the discrimination thresholds evidenced in these experiments. This is the case even if the start/stop latencies are different between causal and non-causal conditions, but fixed within conditions. Our results are therefore only compatible with a change in time perception *during* the interval. In sum then, our results are only compatible with a slower clock rate in causal intervals.

6. General Discussion

The aim of this study was to determine whether temporal binding is effected via the slowing of an internal clock. Specifically, we asked whether distortions in time perception that are typically found in binding preparations are driven by changes in how temporal intervals are perceived.

In Experiments 1 and 2 participants compared causal and non-causal intervals to temporally extended reference stimuli. We estimated the point of subjective equality (PSEs) for a range of durations and conducted regression analyses to determine slope and intercept coefficients. In these experiments we found significantly shallower slopes in causal than non-causal intervals, a clear signature of a slower pacemaker in cause-effect intervals. This mirrors the results of Humphreys and Buehner (2009), who also found shallower slopes for cause-effect intervals using interval estimation. Our results are also comparable with Nolden et al. (2012), who found shorter PSEs for causal, relative to non-causal intervals using the method of constant stimuli. Thus, the results from Experiments 1 and 2 appear to be robust and in line with earlier work in this area. Importantly, the deployment of a range of intervals afforded calculation of regression coefficients, which in turn enabled us to look for evidence pertaining to changes in pacemaker speed and switch latency. Overall, the results from Experiments 1and 2 support a slower clock hypothesis of temporal binding.

In Experiments 3 and 4 we tested the hypothesis that a slower pacemaker during causal episodes would lead to poorer temporal discrimination. Participants were presented either with a sequence of two causal or two non-causal intervals and had to indicate which one was shorter or longer. A difference in pacemaker speed between the conditions should translate into different discrimination thresholds (just noticeable difference: JNDs). In line with our prediction, we found that discrimination thresholds were indeed higher in causal conditions. Furthermore, these results were robust against mere order effects (i.e. the order of reference and comparison intervals in each trial). The results of Experiments 3 and 4 thus corroborate and extend those of Experiments 1 and 2, supporting the hypothesis that intervals delineated by causally related events are timed by a slower clock than intervals delineated by unrelated events.

6.1. Pacemaker slowing or drifts in attention?

It is important to be clear about what the slope coefficient represents. A difference in slopes between causal and non-causal conditions suggests a different number of pulses accumulated between conditions. Given an absence of an intercept difference (which rules out differences in switch latencies), there are two possibilities: one is that there is indeed, a slower pacemaker, while the other is that pacemaker rate is constant, but that there are drifts in attention. The latter suggests that pulses are missed due to non-focal attention, such that certain pulses are not accumulated. Both accounts would explain the shallower slopes found in Experiments 1 and 2. We shall explore each option in turn.

Why would the pacemaker rate decrease in cause-effect intervals? One explanation is that clock rate is linked with arousal. For instance, the (qualitatively) higher predictability associated with causal inference (Sloman, 2005) might result in less arousal during cause-effect intervals. The link between arousal and pacemaker speed is well established (Droit-Volet & Gil, 2016; Droit-Volet & Meck, 2007; Mella, Conty, & Pouthas, 2011). A corollary of this explanation is that clock speed is modulated by cause-effect contingency: if an action always produces a certain outcome with 100% contingency, then causal predictability would be higher than in less contingent preparations. Therefore, it might follow that clock rate decreases as the causal strength between two events increases. One way to test this would be to examine PSE and slope coefficients for different levels of cause-effect contingency. Conversely, replicating Experiments 3 and 4 with varying contingency levels might also produce graded differences in JNDs. Specifically, if clock speed is modulated by cause-effect contingency then JNDs ought to decrease as contingency increases. One caveat is necessary here though: Mere contingency and predictability are insufficient in themselves to affect

clock speed, because non-causal trial events are equally predictable. It is predictability engendered by causality, rather than mere association, that we posit as an explanation for clock slowing (Buehner, 2012).

6.2. Where is attention allocated in temporal binding?

An alternative possibility to clock slowing is that pacemaker rates remain the same in causal and non-causal conditions, but instead, an attentional drift in the former results in fewer pulses accumulated. How might this be achieved? One solution is that more attention is focused on the outcome in causal than non-causal trials, perhaps to confirm that an action did indeed produce the predicted outcome. A simple test of this hypothesis would be to measure the subjective duration of the outcome; more attention would result in a longer perceived outcome durations in causal, compared to non-causal trials. Related to this idea Makwana and Srinivasan (2017) reported that outcomes that were intended by the participants appeared to last subjectively longer than unintended outcomes. Empirically, a systematic drift in attention could also explain the results of Experiments 3 and 4, because fewer clock pulses, overall, would accumulate during causal, than non-causal intervals.

How might we square these time perception approaches with event perception accounts (e.g., Eagleman & Holcombe, 2002)? One possibility is that changes to interval perception influence the perceived time (of occurrence) of events, such that a slower clock between cause and effect would result in the outcome being perceived sooner (and/or the cause being perceived later). In other words, interval timing and event perception might be intertwined, such that pacemaker speed influences the perceived time of occurrence of events. It is important to point out here that this could only be possible if changes in pacemaker speed are specific to the causal interval, rather than

implicating all of temporal perception. Time of occurrence judgments are always made with reference to some other stimulus (e.g. in many temporal binding preparations, participants report the time of their action or its outcome with reference to a fast moving clock hand, cf. Haggard et al., 2002). For a pacemaker account to be able to accommodate shifts in event perception requires that the cause-effect, but not the cause-reference event interval is shortened. In other words, if both intervals were timed by the same pacemaker, both events marking the end of the interval (i.e. the effect and the reference event) would be shifted by the same absolute amount, and thus the difference between, say effect and reference stimulus would remain constant between causal and non-causal preparations. On the other hand, if pacemaker slowing is specific to the cause-effect interval, the cause-reference event interval would not be affected, which would then result in a perceptual shift such that the effect is perceived earlier with reference to the other event (compared to non-causal preparations). We review evidence pertaining to this problem in the next section.

Alternatively, perceived shifts in event perception and a contraction of time in cause-effect intervals might be rooted in two separate processes; causality might independently affect clock speed and the perceived time of events delineating intervals. In other words, time perception might not affect event timing, but instead, the two are distinct processes. Put differently, causality might be the underlying driver for both phenomena, but each is brought about by a different mechanism, each of which is affected by the presence of a causal relation. Either way, the key point is that changes in time perception and perceived event shifts need not necessarily be mutually exclusive.

6.3. General or specific clock slowing?

The results of the current study complement those of recent work from our lab (Fereday and Buehner, 2017), which has addressed the question of whether clock slowing is a general process or unique to the cause-effect interval. In a similar vein to Wenke and Haggard (2009), we used a dual-task paradigm whereby participants either judged the duration of an interval, or a stimulus embedded at different points within the trial. The intervals were random in duration and were delineated by two events: in causal trials, a key press triggered an outcome, and in involuntary trials the disappearance of a visual stimulus was followed by an outcome. Additionally, we scheduled an embedded stimulus to occur before or during the interval, or not at all. We hypothesised that if clock slowing is a general process, its effects would also generalise to embedded events that occur during, but not before, the interval; events embedded into causal intervals would also be judged as shorter than those that occur in non-causal intervals. Naturally, temporal binding should also elicit shorter estimates for causal versus non-causal intervals. In four experiments that factorially varied the modality of embedded events and outcomes (auditory or visual), we found that causal intervals were perceived as shorter than non-causal for all locations of the embedded event (thus replicating the well established binding effect). However, we found no difference between events embedded into causal, from those embedded into non-causal trial intervals. We concluded that internal clock slowing is not a general process, but might instead be unique to cause-effect intervals, such that the rate of a pacemaker timing cause-effect intervals is slower than that tracking the time of other concurrent events. The results of the work reported here now allow us to conclude that these specific changes to interval perception were indeed driven by a slowing of pacemaker speed during causal episodes.

6.4. Does causality or intentional action modulate clock speed?

The literature on temporal binding is divided between the causal and the intentional stance. The latter perspective is unique to the human motor system. According to intentional binding theorists, awareness of voluntary actions and awareness of their sensory consequences are bound together in conscious awareness, reflecting "the general linkage through time between representations of actions and effects" (Haggard et al., 2002 p.384). According to the causal perspective (Buehner & Humphreys, 2009; Buehner, 2012;2015; Eagleman & Holcombe, 2002), temporal binding arises from bi-directional ambiguity reduction in causal inference and time perception. Specifically, temporal contiguity is an important cue to causality (see Buehner, 2005 for an overview): all else being equal, contiguous event sequences are more likely to be judged as causally related than delayed ones. It then follows from Bayesian logic that causally related event sequences are therefore likely to be contiguous in time. Given that time perception is inherently ambiguous, causal event sequences are likely to be perceived as shorter than non-causal sequences. Note that the causal perspective encompasses the intentional one: Intentional binding here is a special subset of causal binding where the cause happens to be an intentional agent.

In the experiments reported here we always operationalised causality via a key press which resulted in an on-screen event and compared this to sequences between two on-screen events. A critical reader might therefore ask to what extent our results genuinely reflect processes driven by causality, or whether they may instead be due to intentionality, or indeed simply the presence of prior motor action? As noted in the Introduction, there is evidence that causality – and not intentional action – is at the root of temporal binding (Buehner, 2012; 2015). Furthermore, Poonian & Cunnington (2013) reported that *self-executed* action-outcome intervals and *observed* action-

outcome intervals lead to comparable binding effects. The mere observation of someone else's causal action is thus sufficient to induce binding – it is not necessary to execute a motor action. Furthermore, Fereday & Buehner's (2017) results (reviewed in the previous section) show that prior motor action in itself (without being linked to a causal relation) is insufficient to induce changes in time perception: Causality is necessary for binding to occur. Given the necessity and sufficiency of causality to produce temporal binding, we are thus confident that the clock-slowing found in our experiments is indeed driven by (perceived) causality.

To conclude, we have demonstrated that temporal causal binding is effected by a slower internal clock. Further experiments are necessary to determine whether clock speed is indeed modulated by the same factors that modulate causal belief.

Acknowledgments

The experiments in this article were conducted as part of RF's PhD thesis at Cardiff University, under the supervision of MJB and SKR.

Data Reference

Fereday, R. (2019, March 7). Temporal binding and internal clocks. Retrieved from osf.io/ar7zk

References

Buehner, M. J. (2005). Contiguity and covariation in human causal inference. *Learning & Behavior 33*(2), 230-238.

Buehner, M. J. (2012). Understanding the past, predicting the future: causation, not intentional action, is the root of temporal binding. *Psychological Science, 23*(12), 1490-1497.

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgment of causality: Effects of prior knowledge, experience, and reinforcement procedure. *Quarterly Journal of Experimental Psychology, 56A,* 865-890.

Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking & Reasoning, 12*(4), 353-378.

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science, 20*(10), 1221-1228.

Cravo, A. M., Haddad, H., Claessens, P. M. E., & Baldo, M. V. C. (2013). Bias and learning in temporal binding: Intervals between actions and outcomes are compressed by prior bias. *Consciousness and Cognition, 22*(4), 1174-1180.

Droit-Volet, S., Fayolle, S., Mathiilde, L., & Gil, S. (2013). Time, emotion and the embodiment of timing. *Timing and time perception, 1(*1), 99-126.

Droit-Volet, S., & Gil, S. (2016). The emotional body and time perception. *Cognition and Emotion, 30*(4), 687-699.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504-513.

Eagleman, D. M., & Holcombe, A. O. (2002). Causality and the perception of time. *Trends in Cognitive Sciences, 6*(8), 323-325.

Engbert, K., Wohlschlager, A., & Haggard, P. (2008). Who is causing what? The sense of agency is relational and efferent-triggered. *Cognition, 107*, 693–704.

Fereday, R., & Buehner, M. J. (2017). Temporal binding and internal clocks: No evidence for general pacemaker slowing. *Journal of Experimental Psychology: Human Perception and Performance, 43*(5), 971-985.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon, & L. G. Allan (Eds), *Annals of the New York Academy of Sciences: Timing and time perception*. New York: New York Academy of Sciences, pp. 52-77.

Haggard, P., & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition, 12*, 695–707.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*(4), 382-385.

Hume, D. (1739/1888). A treatise of human nature. In L. A. Selby-Bigge (Ed.), *Hume's treatise of human nature*. Oxford, UK: Clarendon Press.

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude estimation reveals binding at super-second intervals. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1542-1549.

Kesten, H. (1958). Accelerated stochastic approximation. *Annals of Mathematical Statistics, 29*, 41-59.

Killeen, P. R., & Weiss, N. A. (1987). Optimal timing and the Weber function. *Psychological Review, 94*, 455-468.

Knoblauch, K., & Maloney, L. T. (2012). *Modeling Psychophysical Data in R*. New York: Springer.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain, 106*, 623-642.

Makwana, M & Srinivasan, N (2017) Intended outcome expands in time. *Scientific Reports 7(1).*

Matthews, W. J. (2011). Can we use verbal estimation to dissect the internal clock? Differentiating the effects of pacemaker rate, switch latencies, and judgment processes. *Behavioural Processes, 86*, 68-74.

Matthews, W. J. (2013). How does sequence structure affect the judgment of time? Exploring a weighted sum of segments model. *Cognitive Psychology, 66*, 259-282.

Maxwell, S.E. (2004). The Persistence of Underpowered Studies in Psychological Research: Causes, Consequences and Remedies. *Psychological Methods 9(2)*, 147-163

Mella, N., Conty, L., & Pouthas, V. (2011). The role of physiological arousal in time perception: psychophysiological evidence from an emotion regulation paradigm. *Brain and Cognition, 75*(2), 182-187.

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition, 17*(1), 136-144.

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition, 18*(4), 1056-1064.

Nolden, S., Haering, C., & Kiesel, A. (2012). Assessing intentional binding with the method of constant stimuli. *Consciousness and Cognition, 21*(3), 1176-1185.

Parsons B., Novich, S. D., & Eagleman, D. M. (2013). Motor-sensory recalibration modulates perceived simultaneity of cross-modal events. *Frontiers in Psychology*. 4:46.

Penton-Voak, I. S., Edwards, H., Percival, A., & Wearden, J. H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(3), 307-320.

Pierce, J. (2007). Psychopy – Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.

Rammsayer, T., & Ulrich, R. (2001). Counting models of temporal discrimination. *Psychonomic Bulletin & Review, 8*(2), 270-277.

Ruess, M., Tomaschke, R., & Kiesel, A. (2018). The Time Course of Intentional Binding for Late Effects. *Timing & Time Perception, 6(1)*, 54-70.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgment of causality by human subjects. *Quarterly Journal of Experimental Psychology Section B- Comparative and Physiological Psychology, 41*(2), 139-159.

Shiloh, D., Buehner, M.J. & White, P.A. (2017). The Role of Causality in Temporal Binding: Evidence for an Intentional Boost*. Proceedings of the 39th annual meeting of the cognitive science society*.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). *P*-Curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143*(2), 534-547.

Sloman, S. (2005). *Causal models: How people think about the world and its alternatives*. Oxford: Oxford University Press.

Stetson, C., Fiesta, M. P., & Eagleman, D. M. (2007). Does time really slow down during a frightening event. *PLoS One, 2*(12), e1295.

Tomassini, A, Gori, M, Baud-Bovy, G., Sandini, G. Morrone, M. C, (2014). Motor commands induce time compression for tactile stimuli. *Journal of Neuroscience 34(27), 9164-9172.*

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research, 35*(17), 2503-2522.

Tsakiris, M., & Haggard, P. (2003). Awareness of somatic events associated with a voluntary action. *Experimental Brain Research, 149*, 439–446.

Tse, P. U., Rivest, J., Intriligator, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics, 6*6(7), 1171-1189.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology*, *51B*, 97-120.

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: conformity and violations. *Quarterly Journal of Experimental Psychology, 61*(4), 569-587.

Wearden, J. H., Norton, R., Martin, S., Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 716-729.

Wenke, D., & Haggard, P. (2009). How voluntary actions modulate time perception. *Experimental Brain Research, 196*(3), 311-318.

Footnotes

1 We also calculated Just-Noticeable-Differences (JNDs) and regressed them over standard durations to check for conformity with scalar assumptions (Wearden & Lejeune, 2008). This can be found in the online supplement.

2 We used 3 standard deviations as an arbitrary rule to apply uniformly in all experiments. Note that no participants were removed in Experiments 3 and 4. We attribute this to the stimuli used: In Experiments 1 and 2, the two intervals were delineated by separate events (two events demarked interval one, and a single event marked interval 2). In Experiments 3 and 4, both intervals were delineated by two events. We initially assumed that the latter design would confuse participants as to which intervals required discriminating. We note that the increased noise in Experiments 1 and 2 likely reflects the difficulty of the task (relative to Experiments 3 and 4). Importantly, we further note that our use of the stimuli in this manner merely adds noise – it does not affect credibility of the results; the same pattern of results would be expected if less noise were present.

3 Note that Figure 2 displays the mean regression lines, whereas the coefficients here were derived from regressions fitted to individual participants.